Perspectives on Human Error: Hindsight Biases and Local Rationality

David D. Woods
Cognitive Systems Engineering Laboratory
Institute for Ergonomics
The Ohio State University

Richard I. Cook
Department of Anesthesia and Critical Care
University of Chicago

Introduction

Early Episodes of "Human Error"

Consider these episodes where some stakeholders reacted to failure by attributing the cause to "human error," but where more careful examination showed how a combnation of factors created the conditions for failure.[1]

Case 1:

- In 1796 the astronomer Maskelyne fired his assistant Kinnebrook because the latter's observations did not match his own (see Boring, 1950).

- Bessel, another astronomer, studied the case empirically and identified systematic factors[2] which produced imprecise observations.

[1] The term stakeholders refers to all of the different groups who are affected by an accident in that domain and by changes to operations, regulations, equipment, policies, etc. as a result of reactions to that accident. For example, an accident in aviation affects the public as consumers of the service, the FAA as regulators of the industry, pilots as practitioners, air traffic controllers as practitioners, air carrier organizations as providers of the service, manufacturers of the aircraft type as equipment designers, other development organizations that develop other equipment such as avionics, navigation aids, and software, and other groups as well.

[2] By systematic factors, we mean the behavior is not random but lawful, i.e., there are empirical regularities, factors that influence behavior are external to individuals (system properties), these factors have an effect because they influence the physical, cognitive and collaborative activities of practitioners, and, finally, because there are regularities, there are predictable effects.

The implicit assumption was that one person (the assistant) was the source of failure whether due to some inherent trait or to lack of effort on his part. Bessel broke free of this assumption and empirically examined individual differences in astronomical observations. He found that there were wide differences across observers given the methods of the day. The techniques for making observations at this time required a combination of auditory and visual judgments. Those judgments were shaped by the tools of the day, pendulum clocks and telescope hairlines, in relation to the demands of the task. Dismissing Kinnebrook did not change what made the task difficult, did not eliminate individual differences, and did not make the task less vulnerable to sources of imprecision. Progress was based on searching for better methods for making astronomical observations, re-designing the tools that supported astronomers, and re-designing the tasks to change the demands placed on human judgment.

Case 2:

- In 1947 investigations of military aviation accidents concluded that pilot errors were the cause of the crashes.

- Fitts and Jones empirically studied pilot performance in the cockpit and showed how systematic factors in interpreting instruments and operating controls produced misassessments and actions not as intended (see Fitts and Jones, 1947).

The implicit assumption was that the person closest to the failure was the cause. Investigators saw that the aircraft was in principle flyable and that other pilots were able to fly such aircraft successfully. They could show how the necessary data were available for the pilot to correctly identify the actual situation and act in an appropriate way. Since the pilot was the human closest to the accident who could have acted differently, it seemed obvious to conclude that the pilot was the cause of the failure.

Fitts and his colleague empirically looked for factors that could have influenced the performance of the pilots. They found that, given the design of the displays and layout of the controls, people relatively often misread instruments or operated the wrong control, especially when task demands were high. The misreadings and misoperations were design-induced in the sense that researchers could link properties of interface design to these erroneous actions and assessments. In other words, the “errors” were not random events, rather they resulted from understandable, regular, and predictable aspects of the design of the tools practitioners used.

The researchers found that misreadings and misoperations occurred, but did not always lead to accidents due to two factors. First, pilots often detected these errors before negative consequences occurred. Second, the misreadings and misoperations alone did not lead directly to an accident. Disaster or near misses usually occurred only when these errors occurred in combination with other factors or other circumstances.

In the end, the constructive solution was not to conclude that pilots err, but rather to understand principles and techniques for the design of visual displays and control layout. Changing the artifacts used by pilots changed the demands on human perception and cognition and changed the performance of pilots.

Erratic People or System Factors?

While historical, the above episodes encapsulate current, widespread beliefs in many technical and professional communities and the public in general about the nature of human error and how systems fail. In most domains today, from aviation to industrial processes to transportation systems to medicine, when systems fail we find the same pattern as was observed in these earlier cases.

1. Stakeholders claim failure is "caused" by unreliable or erratic performance of individuals working at the sharp end[3] who undermine systems which otherwise worked as designed (for example, see the recent history of pilot-automation accidents in aviation; Billings, 1996; Woods and Sarter, in press). The search for causes tends to stop when we can find the human or group closest to the accident who could have acted differently in a way that would have led to a different outcome. These people are seen as the source or "cause" of the failure, that is, the outcome was due to "human error." When stakeholders see erratic people as the cause of bad outcomes, then they respond by calls to remove these people from practice, to provide remedial training to other practitioners, to urge other practitioners to try harder, and to regiment practice through policies, procedures, and automation.

2. However, researchers look more closely at the system in which these practitioners are embedded and their studies reveal a different picture.[4] Their results show how popular beliefs that such accidents are due simply to

[3] It has proven useful to depict complex systems such as health care, aviation and electrical power generation and others as having a sharp and a blunt end (Reason, 1990). At the sharp end, practitioners interact with the underlying process in their roles as pilots, spacecraft controllers, and, in medicine, as nurses, physicians, technicians, and pharmacists. At the blunt end of a system are regulators, administrators, economic policy makers, and technology suppliers. The blunt end of the system controls the resources and constraints that confront the practitioner at the sharp end, shaping and presenting sometimes conflicting incentives and demands (Reason, in press).

[4] The term practitioner refers to "a person engaged in the practice of a profession or occupation" (Webster's, 1990).

isolated blunders of individuals mask the deeper story -- a story of multiple contributors that create the conditions that lead to operator errors. Reason (1990, p. 173) summarizes the results: “Rather than being the main instigators of an accident, operators tend to be the inheritors of system defects... Their part is that of adding the final garnish to a lethal brew whose ingredients have already been long in the cooking.” The empirical results reveal regularities in organizational dynamics and in the design of artifacts that produce the potential for certain kinds of erroneous actions and assessments by people working at the sharp end of the system (Woods, Johannesen, Cook and Sarter, 1994; Reason, 1998).

For example, one basic finding from research on disasters in complex systems (e.g., Reason, 1990) is that accidents are not due to a single failure or cause. Accidents in complex systems only occur through the concatenation of multiple small factors or failures, each necessary but only jointly sufficient to produce the accident. Often, these small failures or vulnerabilities are present in the organization or operational system long before a specific incident is triggered. All complex systems contain such “latent” factors or failures, but only rarely do they combine to create an accident (Reason, 1998).

This pattern of multiple, latent factors occurs because the people in an industry recognize the existence of various hazards that threaten to cause accidents or other significant consequences, and they design defenses that include technical, human, and organizational elements. For example, people in health care recognize the hazards associated with the need to deliver multiple drugs to multiple people at unpredictable times in a hospital setting and use computers, labeling methods, patient identification cross-checking, staff training, and other methods to defend against misadministrations. Accidents in these kinds of systems occur when multiple factors join together to create the trajectory for an accident by eroding, bypassing or breaking through the multiple defenses. Because there are a set of contributors, multiple opportunities arise to redirect the trajectory away from disaster. The research has revealed that an important part of safety is enhancing opportunities for people to recognize that a trajectory is heading closer to a poor outcome and to recover before negative consequences occur (Rasmussen, 1986). Factors that reduce error tolerance or block error detection and recovery degrade system performance.

Behind the Label Human Error

Based on this pattern in the data, researchers see that the label “human error” should serve as the starting point for investigating how systems fail, not as a conclusion. In other words, human performance is shaped by systematic factors, and the scientific study of failure is concerned with understanding how these factors shape the cognition, collaboration and ultimately the behavior of people in various work domains.

This research base has identified some of these regularities. In particular, we know about how a variety of factors make certain kinds of erroneous actions and assessments predictable (e.g., Norman, 1983; Norman, 1988; Hollnagel, 1993). Our ability to predict the timing and number of erroneous actions is very weak, but our ability to predict the kind of errors that will occur, when people do err, is often good. For example, when research pursues this deeper story behind the label "human error," we find imbalances between the demands practitioners face and the resources available to meet the demands of that field of activity (Rasmussen, 1986). These demand-resource imbalances can affect the development of necessary expertise (Feltovich, Ford, and Hoffman, 1997), how the system brings additional expertise to bear especially when more difficult problems emerge, how people cope with multiple pressures and demands (Klein, 1998), how the system supports cooperative work activities especially when the tempo of operations increases, and how organizational constraints hinder or aid practitioners when they face difficult tradeoffs and dilemmas (Weick and Roberts, 1993).

This chapter will explore only a small portion of the issues that come to the fore when one goes behind the label human error. The space of social, psychological, technological, and organizational issues is large in part because people play such diverse roles in work environments and because work environments themselves vary so much.

Two Perspectives: Studying the Factors that Affect Human Performance and Studying Reactions to Failure

The chapter revolves around an ambiguity in the label human error. The two historical episodes introduced at the beginning of the chapter illustrate this ambiguity.

When we use the label human error we are sometimes referring to the processes and factors that influence the behavior of the people in the situation. From this perspective investigators are trying to understand the factors that lead up to erroneous actions and assessments. In this sense, the label human error points at all of the factors that influence human performance -- in particular, the performance of the practitioners working in some field of activity. In the examples at the beginning of the chapter, Bessel and Fitts followed up the incidents with this kind of investigation. Understanding human performance is a very large subject (in part the subject of much of this volume) and here we will examine only a few of the relevant issues -- predominately cognitive factors, but also how those cognitive processes are shaped by artifacts, coordination across multiple people, and organizational pressures.

While the above constitutes the bulk of this chapter, the label human error can refer to a different class of psychological issues and phenomena. After-the-fact, stakeholders look back and make judgments about what led to the accident or incident. In the examples at the beginning of the chapter, Maskelyne and the authors of the aviation accident reports reacted after-the-fact with the judgment that individuals were the cause of the accidents. Labeling a past action as erroneous is a judgment based on a different perspective and on different information than what was available to the practitioners in context. In other words, this judgment is a process of causal attribution, and there is an extensive body of research about the social and psychological factors which influence these kinds of attributions of causality (e.g., Kelley, 1973; Fischhoff, 1975; Baron and Hershey, 1988; Hilton, 1990). From this perspective, error research studies the social and psychological processes which govern our reactions to failure as stakeholders in the system in question (Tasca, 1990; Woods et al., 1994, chapter 6).

Our reactions to failure as stakeholders are influenced by many factors. One of the most critical is that, after an accident, we know the outcome and, working backwards, what were critical assessments or actions that, if they had been different, would have avoided that outcome. It is easy for us with the benefit of hindsight to say, "how could they have missed x?" or "how could they have not realized that x would obviously lead to y."

Studies have consistently shown that people have a tendency to judge the quality of a process by its outcome (Baron and Hershey, 1988; ; Lipshitz, 1989; Caplan, Posner and Cheney, 1991). In a typical study, two groups are asked to evaluate human performance in cases with the same descriptive facts but with the outcomes randomly assigned to be either bad or neutral. Those with knowledge of a poor outcome judge the same decision or action more severely. This is referred to as the outcome bias (Baron and Hershey, 1988) and has been demonstrated with practitioners in different domains. For example, Caplan, Posner, and Cheney (1991) found an inverse relationship between the severity of outcome and anesthesiologists' judgments of the appropriateness of care. The judges consistently rated the care in cases with bad outcomes as substandard while viewing the same behaviors with neutral outcomes as being up to standard even though the care (i.e., the preceding human performance) was identical. The information about outcome biased the evaluation of the process that was followed.

Other research has shown that once people have knowledge of an outcome, they tend to view the outcome as having been more probable than other possible outcomes. Moreover, people tend to be largely unaware of the modifying effect of outcome information on what they believe they could have known in foresight. These two tendencies collectively have been termed the hindsight bias. Fischhoff (1975) originally demonstrated the hindsight bias in a set of experiments that compared foresight and hindsight judgments concerning the likelihood of particular socio-historical events. Basically, the bias has been demonstrated in the following way: participants are told about some event, and some are provided with outcome information. At least two different outcomes are used in order to control for one particular outcome being a priori more likely. Participants are then asked to estimate the probabilities associated with the several possible outcomes. Participants given the outcome information are told to ignore it in coming up with their estimates, i.e., "to respond as if they had not known the actual outcome," or in some cases are told to respond as they think others without outcome knowledge would respond. Those participants with the outcome knowledge judge the outcomes they had knowledge about as more likely than the participants without the outcome knowledge, even when those making the judgments have been warned about the phenomenon and been advised to guard against it (Fischhoff, 1982). Experiments on the hindsight bias have shown that: (a) people overestimate what they would have known in foresight, (b) they also overestimate what others knew in foresight, and (c) they actually misremember what they themselves knew in foresight.

Taken together, the outcome and hindsight biases have strong implications for error analyses.

- Decisions and actions having a negative outcome will be judged more harshly than if the same process had resulted in a neutral or positive outcome. We can expect this result even when judges are warned about the phenomenon and have been advised to guard against it.
- Judges will tend to believe that people involved in some incident knew more about their situation than they actually did. Judges will tend to think that people should have seen how their actions would lead up to the outcome failure.

One sense of studying "human error" then involves understanding how social and psychological processes such as hindsight and outcome biases shape our reactions to failure as stakeholders in the failed system.[5] In general, we react, after the fact, as if the knowledge we now possess was available to the operators then. This oversimplifies or trivializes the situation confronting

[5] In a narrow sense, the outcome and hindsight biases refer to specific experimental findings from different test paradigms. In a broader sense, both of these experimental results, and other results, refer to a collection of factors that influence our reactions to failures based on information that is available only after the outcome is known. In the context of error analysis, we have used the label "hindsight bias" to refer to the broader perspective of a judge looking back in hindsight to evaluate the performance of others (Woods et al., 1994). Both specific experimental results illustrate ways in which people with the benefit of hindsight can misperceive and misanalyze the factors that influenced the behavior of the people working in the situation before outcome was known.

the practitioners, and masks the processes affecting practitioner behavior before-the-fact. As a result, hindsight and outcome bias blocks our ability to see the deeper story of systematic factors that predictably shape human performance.

There is limited data available about the reactions of stakeholders to failure (but see, Tasca, 1990 and the review in Woods et al., 1994, chapter 6). This perspective shifts the focus of investigation from studying sharp end practitioners to studying stakeholders. It emphasizes the need to collect data on reactions to failure and to contrast those reactions after-the-fact to results on the factors that influence human performance before-the-fact derived from methods that reduce hindsight biases. Such studies can draw from the conceptual base created by experimental studies on the social and psychological factors in judgments of causal attribution such as hindsight biases.

The main portion of this chapter will focus on the deeper story behind the label human error of factors that influence human performance, especially some of the cognitive factors. Then, we will return to hindsight and outcome bias to illustrate several misconceptions about cognition that arise often when incidents are reviewed with knowledge of outcome.

Cognitive Factors and Human Performance

What cognitive factors affect the performance of practitioners in complex settings like medicine, aviation, telecommunications, process plants, and space mission control? There are many ways one could organize classes of cognitive factors relevant to human performance (e.g., Rasmussen, 1986; Norman, 1988; Reason, 1990). Different aspects of cognition will be relevant to various settings and situations. For example, one area of human performance is concerned with slips of action such as capture errors or omissions of isolated acts (e.g., Norman, 1981; Reason and Mycielska, 1982; Byrne and Bovair, 1997). We have found it useful to teach people about cognitive factors and error by using the concept of bounded or local rationality (Simon, 1957).

Local Rationality

At work, groups of practitioners pursue goals and match procedures to situations, but they also

- resolve conflicts,
- anticipate hazards,
- accommodate variation and change,

- cope with surprise,
- workaround obstacles,
- close gaps between plans and real situations,
- detect and recover from miscommunications and misassessments.

In these activities practitioners at the sharp end block potential accident trajectories. In other words, people actively contribute to safety when they can carry out these roles successfully. “Error” research on human performance tries to identify factors that undermine practitioners’ ability to do these activities successfully. The question then becomes how can the same processes result in success some of the time but result in failure in other circumstances (Rasmussen, 1986).

The concept of bounded rationality is very useful for helping us think about how people can forms intentions act in ways that later events will reveal are erroneous. Peoples’ behavior in the work situations can be considered as consistent with Newell’s principle of rationality—that is, practitioners use their knowledge to pursue their goals (Newell, 1982). But there are bounds to the data that they pick up or search out, bounds to the knowledge that they possess, bounds to the knowledge that they activate in a particular context, and there may be multiple goals which conflict (Simon, 1957). In other words, people's behavior can be seen as “rational,” though possibly erroneous, when seen from the point of view of their knowledge, their mindset, and the multiple goals they are trying to balance (Rasmussen, Duncan and Leplat, 1987). Rationality here does not mean consistent with external, global standards such as models, policies or procedures; rationality in Newell’s principle is defined locally from the point of view of the people in a situation as they use their knowledge to pursue their goals based on their view of the situation. As a result, for the context of error, we will refer to the concept that human rationality is limited or bounded as “local” rationality (Woods et al., 1994).

Fundamentally, human (and real machine) problem-solvers possess finite capabilities. They cannot anticipate and consider all the possible alternatives and information that may be relevant in complex problems. This means that the rationality of finite resource problem solvers is local in the sense that it is exercised relative to the complexity of the environment in which they function (Klein et al., 1993; Klein, 1998). It takes effort (which consumes limited computational resources) to seek out evidence, to interpret it (as relevant), and to assimilate it with other evidence. Evidence may come in over time, over many noisy channels. The process may yield information only in response to diagnostic interventions. Time pressure, which compels action (or the de facto decision not to act), makes it impossible to wait for all evidence to accrue. Multiple goals may be relevant, not all of which are consistent. It may not be clear, in foresight, which goals are the most important ones to focus on at any one particular moment in time. Human problem solvers cannot handle all the potentially relevant information, cannot activate and hold in mind all of the relevant knowledge, and cannot entertain all potentially relevant trains of thought. Hence, rationality must be local—attending to only a subset of the possible evidence or knowledge that could be, in principle, relevant to the problem.

The role for "error" research, in the sense of understanding the factors that influence human performance, is to understand how limited knowledge (missing knowledge or misconceptions), how a limited and changing mindset, and how multiple interacting goals shaped the behavior of the people in the evolving situation. In other words, this type of error research reconstructs what the view was like or would have been like had we stood in the same situation as the participants. If we can understand how their knowledge, their mindset, and their goals guided the behavior of the participants, then we can see how they were vulnerable to err given the demands of the situation they faced. We can see new ways to help practitioners activate relevant knowledge, shift attention to the critical focus among multiple tasks in a rich, changing data field, and recognize and balance competing goals.

Given that people use their knowledge to pursue their goals, but given that there are bounds to their knowledge, limits to their mindset and multiple not always consistent goals to achieve, one can learn about the performance of practitioners at the sharp end by looking at factors that affect

- how knowledge relevant to the situation at hand is called to mind -- knowledge in context,
- how we come to focus on one perspective on or one part of a rich and changing environment and how we shift that focus across multiple events over time -- mindset,
- how we balance or make tradeoffs among multiple interacting goals -- interacting goals.

We have found it useful to group various findings and concepts into these three classes of cognitive factors that govern how people form intentions to act (Cook and Woods, 1994). Problems in the coordination of these cognitive functions, relative to the demands imposed by the field of activity, create the potential for mismanaging systems towards failure. For example, in terms of knowledge factors, some of the possible problems are buggy knowledge (e.g., incorrect model of device function), inert knowledge, and oversimplifications (Spiro, Coulson, Feltovich, and Anderson, 1988). In terms of mindset, one form of breakdown occurs when an inappropriate mindset takes hold or persists in the face of evidence which does not fit this assessment. Failures very often can be traced back to dilemmas and tradeoffs that arise from multiple interacting and sometimes conflicting goals. Practitioners by the very nature of their role at the sharp end of systems must implicitly or explicitly resolve these conflicts and dilemmas as they are expressed in particular situations (Cook and Woods, 1994).

Knowledge in Context

Knowledge factors refer to the process of bringing knowledge to bear to solve problems in context:

- what knowledge do practitioners possess about the system or process in question (is it correct, incomplete, or erroneous, i.e., "buggy"?),
- how this knowledge is organized so that it can be used flexibly in different contexts, and
- the processes involved in calling to mind the knowledge relevant to the situation at hand.

Knowledge of the world and its operation may be complete or incomplete and accurate or inaccurate. Practitioners may act based on inaccurate knowledge or on incomplete knowledge about some aspect of the complex system or its operation. When the mental model that practitioners hold of such systems is inaccurate or incomplete, their actions may well be inappropriate. These mental models are sometimes described as "buggy." The study of practitioners' mental models has examined the models that people use for understanding technological, physical, and physiological processes. Several volumes are available which provide a comprehensive view of research on this question (see Gentner and Stevens, 1983; Chi, Glaser, and Farr, 1988; Feltovich, Ford and Hoffman, 1997).

Note that research in this area has emphasized that mere possession of knowledge is not enough for expertise. It is also critical for knowledge to be organized so that it can be activated and used in different contexts (Bransford, Sherwood, Vye, and Rieser, 1986). Thus, Feltovich, Spiro, and Coulson (1989) and others emphasize that one component of human expertise is the <u>flexible</u> application of knowledge in <u>new</u> situations.

There are multiple overlapping lines of research related to the activation of knowledge in context by humans performing in complex systems. These include:

- the problem of inert knowledge, and
- the use of heuristics, simplifications, and approximations.

Going behind the label "human error" involves investigating how knowledge was or could have been brought to bear in the evolving incident. Any of the above factors could influence the activation of knowledge in context—for example, did the participants have incomplete or erroneous knowledge? Were otherwise useful simplifications applied in circumstances that demanded consideration of a deeper model of the factors at work in the case? Did relevant knowledge remain inert? We will briefly sample a few of the issues in this area.

Activating Relevant Knowledge In Context: The Problem Of Inert Knowledge

Lack of knowledge or buggy knowledge may be one part of the puzzle, but the more critical question may be factors that affect whether relevant knowledge is activated and utilized in the actual problem solving context (e.g., Bransford et al., 1986). The question is not just does the problem-solver know some particular piece of domain knowledge, but does he or she call it to mind when it is relevant to the problem at hand and does he or she know how to utilize this knowledge in problem solving? We tend to assume that if a person can be shown to possess a piece of knowledge in one situation and context, then this knowledge should be accessible under all conditions where it might be useful. In contrast, a variety of research results have revealed dissociation effects where knowledge accessed in one context remains inert in another (Gentner and Stevens, 1983; Perkins and Martin, 1986).

Thus, the fact that people possess relevant knowledge does not guarantee that this knowledge will be activated when needed. The critical question is not to show that the problem-solver possesses domain knowledge, but rather the more stringent criterion that situation-relevant knowledge is accessible under the conditions in which the task is performed. Knowledge that is accessed only in a restricted set of contexts is called inert knowledge. Inert knowledge may be related to cases that are difficult to handle, not because problem-solvers do not know the individual pieces of knowledge needed to build a solution, but because they have not confronted previously the need to join the pieces together.

Results from accident investigations often show that the people involved did not call to mind all the relevant knowledge during the incident although they “knew” and recognized the significance of the knowledge afterwards. The triggering of a knowledge item *X* may depend on subtle pattern recognition factors that are not present in every case where *X* is relevant. Alternatively, that triggering may depend critically on having sufficient time to process all the available stimuli in order to extract the pattern. This may explain the difficulty practitioners have in “seeing” the relevant details when the pace of activity is high.

One implication of these results is that training experiences should exercise knowledge in the contexts where it is likely to be needed.

Oversimplifications

People tend to cope with complexity through simplifying heuristics. Heuristics are useful because they are usually relatively easy to apply and minimize the cognitive effort required to produce decisions. These simplifications may be useful approximations that allow limited resource practitioners to function robustly over a variety of problem demand factors or they may be distortions or misconceptions that appear to work satisfactorily under some conditions but lead to error in others. Feltovich et al. (1989) call the latter "oversimplifications."

In studying the acquisition and representation of complex concepts in biomedicine, Feltovich et al. (1989) found that various oversimplifications were held by some medical students and even by some practicing physicians. They found that ". . . bits and pieces of knowledge, in themselves sometimes correct, sometimes partly wrong in aspects, or sometimes absent in critical places, interact with each other to create large-scale and robust misconceptions" (Feltovich et al., 1989, p. 162). Examples of kinds of oversimplification include (see Feltovich, Spiro, and Coulson, 1993):

- seeing different entities as more similar than they actually are,
- treating dynamic phenomena as static,
- assuming that some general principle accounts for all of a phenomenon,
- treating multidimensional phenomena as unidimensional or according to a subset of the dimensions,
- treating continuous variables as discrete,
- treating highly interconnected concepts as separable,
- treating the whole as merely the sum of its parts.

Feltovich and his colleagues' work has important implications for the teaching and training of complex material. Their studies and analyses challenge the view of instruction that presents initially simplified material in modules that decompose complex concepts into their simpler components with the belief that these will eventually "add up" for the advanced learner (Feltovich et al., 1993). Instructional analogies, while serving to convey certain aspects of a complex phenomenon, may miss some crucial ones and mislead on others. The analytic decomposition misrepresents concepts that have interactions among variables. The conventional approach may produce a false sense of understanding and inhibit pursuit of deeper understanding because learners may resist learning a more complex model once they already have an apparently useful simpler one (Spiro et al., 1988). Feltovich and his colleagues have developed the theoretical basis for a new approach to advanced knowledge acquisition in ill-structured domains (Feltovich, Spiro, and Coulson, 1997).

Why do practitioners utilize simplified or oversimplified knowledge? These simplifying tendencies may occur because of the cognitive effort required in demanding circumstances (Feltovich et al., 1989). Also, simplifications may be adaptive, first, because the effort required to follow more "ideal" reasoning paths may be so large that it would keep practitioners from acting with the speed demanded in actual environments. This has been shown elegantly by Payne, Bettman, and Johnson (1988) and by Payne, Johnson, Bettman, and

Coupey (1990) who demonstrated that simplified methods will produce a higher proportion of correct choices between multiple alternatives under conditions of time pressure.

In summary, heuristics represent effective and necessary adaptations to the demands of real workplaces (Rasmussen, 1986). The issue may not always be the shortcut or simplification itself, but whether practitioners know the limits of the shortcuts, can recognize situations where the simplification is no longer relevant, and have the ability to use more complex concepts, methods, or models (or the ability to integrate help from specialist knowledge sources) when the situation they face demands it

Mindset

"Everyone knows what attention is. It is the taking possession by the mind, in a clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought" (James, 1890 I, 403-404). We are able to focus, temporarily, on some objects, events, actions in the world or on some of our goals, expectations or trains of thought while remaining sensitive to new objects or events that may occur. We can refer to this broadly as a state of attentional focus or mindset (LaBerge, 1995).

Mindset is not fixed, but shifts to explore the world and to track potentially relevant changes in the world (LaBerge, 1995). In other words, one re-orients attentional focus to a newly relevant object or event from a previous state where attention was focused on other objects or on other cognitive activities (such as diagnostic search, response planning, communication to others). New stimuli are occurring constantly; any of these could serve as a signal we should interrupt ongoing lines of thought and re-orient attention. This re-orientation involves disengagement from a previous focus and movement of attention to a new focus. Interestingly, this control of attentional focus can be seen as a skillful activity that can be developed through training (Gopher, 1991) or supported (or undermined) by the design of artifacts and intelligent machine agents (Woods 1995; Patterson, Watts-Perotti, and Woods, in press).

A basic challenge for practitioners at work is where to focus attention next in a changing world (Woods and Watts, 1997). Which object, event, goal or line of thought we focus on depends on the interaction of two sets of activity. One of these is goal or knowledge directed, endogenous processes (often called attentional set) that depend on the observer's current knowledge, goals and expectations about the task at hand. The other set of processes are stimulus- or data-driven where attributes of the stimulus world (unique features, transients, new objects) elicit attentional capture independent of the observer's current mindset (Yantis, 1993). These salient changes in the world help guide shifts focus of attention or mindset to relevant new events, objects, or tasks.

These two kinds of processes combine in a cycle, what Neisser (1976) called the perceptual cycle, where unique events in the environment shift the focus of attention or mindset, call to mind knowledge, trigger new lines of thought. The activated knowledge, expectations or goals in turn guides further exploration and action. This cycle is a crucial concept for those trying to understand human performance in the workplace (e.g., Jager Adams, Tenney, and Pew, 1995).

There are a variety of interesting psychological phenomena that can be organized under the label "mindset." Examples include:

- attentional control and loss of situation awareness (Gopher, 1991; Jager Adams et al., 1995: Durso and Gronlund, this volume),
- revising assessments as new evidence occurs, for example in garden path problems and in failures to revise, such as fixation effects (Johnson et al., 1988; DeKeyser and Woods, 1990),
- framing effects and the representation effect (Johnson et al., 1991; Zhang and Norman, 1994; Zhang, 1997),
- juggling multiple lines of thought and activity in time including breakdowns in workload management and thematic vagabonding (Dorner, 1983).

There are a variety of problems that can occur in synchronizing mindset to goals and priorities in a changing world depending on problem demands, skill levels, coordinative structures, and the artifacts available to support performance, e.g.:

- Irrelevant stimuli may intrude on a primary task, e.g., distraction -- a breakdown in selective attention.
- One's mindset may be too easily interrupted, as in thematic vagabonding[6] -- a breakdown in attention switching.
- One's mindset may be too hard to interrupt and re-focus, i.e., fixating on one view of the problem -- a breakdown in attention switching.
- Attention may be captured by irrelevant stimuli or relevant stimuli fail to capture attention in a cognitively noisy workplace, e.g., habituation to nuisance alarms or 'high rates of false alarms (Woods, 1995; Getty, Swets, Pickett, and Gonthier, 1995).

[6] Thematic vagabonding refers to one form of loss of coherence where multiple interacting themes are treated superficially and independently so that the person or team jumps incoherently from one theme to the next (Dorner, 1983).

- Breakdowns can occur in setting priorities or making tradeoffs, e.g., shedding the wrong task under high workload -- a breakdown in attention switching.

When the goal of the investigator is to understand the factors that shape human performance, a critical step is to trace out how an individual's mindset or a group's mindset develops and changes as the incident evolves. Often this involves understanding how the mindset of different individuals or groups interact and or remain encapsulated. This analysis reveals the cues which attracted attention, how did they match the expectations of the observers, what was called to mind, what lines of thought were triggered, and how those goals and issues guided further exploration and action, which in turn generates new cues. This analysis reveals the kinds of attentional challenges produced by the evolving problem and the kinds of attentional breakdowns that occurred.

Loss of Situation Awareness

Situation awareness is a label that is often used to refer to many of the cognitive processes involved in forming and changing mindset (e.g., Jager Adams et al., 1995; Durso and Gronlund, this volume). Maintaining situation awareness necessarily requires shifts of attention to inform and modify a coherent picture or model of the situation the practitioners face. Anticipating how the situation may develop or change in the future state is a particularly important aspect.

Breakdowns in these cognitive processes can lead to operational difficulties in handling the demands of dynamic, event-driven incidents. In aviation circles this is known as "falling behind the plane" and in aircraft carrier flight operations it has been described as "losing the bubble" (Roberts and Rousseau, 1989). In each case what is being lost is the operator's internal representation of the state of the world at that moment and the direction in which the forces active in the world are taking the system that the operator is trying to control.

Fischer, Orasanu and Montvalo (1993) examined the juggling of multiple threads of a problem in a simulated aviation scenario. More effective crews were better able to coordinate their activities with multiple issues over time; less effective crews traded one problem for another. More effective crews were sensitive to the interactions between multiple threads involved in the incident; less effective crews tended to simplify the situations they faced and were less sensitive to the constraints of the particular context they faced. Less effective crews "were controlled by the task demands" and did not look ahead or prepare for what would come next. As a result, they were more likely to run out of time or encounter other cascading problems. Interestingly, there were written procedures for each of the problems the crews faced. The cognitive work associated with managing multiple threads of activity goes beyond the activities needed to merely follow the rules.

This study illustrates how breakdowns in attentional control result from how task demands challenge human abilities to shift and focus attention in a changing task world (Gilson, 1995).

Failures To Revise Situation Assessments: Fixation Or Cognitive Lockup
Diagnostic problems fraught with inherent uncertainties are common, especially when evidence arrives over time and situations can change. Incidents rarely spring full blown and complete; incidents evolve. Practitioners make provisional assessments and form expectancies based on partial and uncertain data. These assessments are incrementally updated and revised as more evidence comes in. Furthermore, situation assessment and plan formulation are not distinct sequential stages, but rather they are closely interwoven processes with partial and provisional plan development and feedback leading to revised situation assessments (Woods, 1994).

As a result, it may be necessary for practitioners to entertain and evaluate what turn out later to be erroneous assessments. Problems arise when the revision process breaks down and the practitioner becomes fixated on an erroneous assessment, missing, discounting or re-interpreting discrepant evidence (e.g., DeKeyser and Woods, 1990; Johnson et al., 1981, 1988; Gaba and DeAnda, 1989). These failures to revise situation assessment as new evidence comes in have been referred to as functional fixations, cognitive lockup and cognitive hysteresis. The operational teams involved in several major accidents (e.g., the Three Mile Island accident) seem to have exhibited this pattern of behavior (Woods et al., 1987).

In cases of fixation, the initial situation assessment tends to be appropriate, in the sense of being consistent with the partial information available at that early stage of the incident. As the incident evolves, however, people fail to revise their assessments in response to new evidence, evidence that indicates an evolution away from the expected path. The practitioners become fixated on the old assessment and fail to revise their situation assessment and plans in a manner appropriate to the data now present in their world. Thus, a fixation occurs when practitioners fail to revise their situation assessment or course of action and maintain an inappropriate judgment or action in the face of opportunities to revise. Thus, fixations represent breakdowns in the process of error detection and recovery where people discount discrepant evidence and fail to keep up with new evidence or a changing situation.

Several criteria need to be met in order to describe an event as a fixation. One critical feature is that there is some form of persistence over time in the behavior of the fixated person or team. Second, opportunities to revise are cues, available or potentially available to the practitioners, that could have started the revision process if observed and interpreted properly. In part, this feature distinguishes fixations from simple cases of inexperience, lack of knowledge, or other problems that impair error detection and recovery. The basic defining characteristic of fixations is that the immediate problem-solving context has biased the practitioners' mindset in some direction inappropriately. In naturally occurring problems, the context in which the incident occurs and the way the incident evolves activates certain kinds of knowledge as relevant to the evolving incident. This knowledge, in turn, affects how new incoming information is interpreted -- a framing effect. After the fact or after the correct diagnosis has been pointed out, the solution seems obvious, even to the fixated person or team.

There are certain types of problems that may encourage fixations by mimicking other situations, in effect, leading practitioners down a garden path (Johnson et al., 1988; Johnson, Jamal, and Berryman, 1991; Johnson, Grazioli, Jamal, and Zualkernan, 1992; Roth, Woods, and Pople, 1992). In garden path problems "early cues strongly suggest [plausible but] incorrect answers, and later, usually weaker cues suggest answers that are correct" (Johnson, Moen and Thompson, 1988). It is important to point out that the erroneous assessments resulting from being led down the garden path are not due to knowledge factors. Rather, they seem to occur because "a problem-solving process that works most of the time is applied to a class of problems for which it is not well suited" (Johnson et al., 1988). This notion of garden path situations is important because it identifies a task genotype (Hollnagel, 1993) in which people become susceptible to fixations. The problems that occur are best attributed to the interaction of particular environmental (task) features and the heuristics people apply (local rationality given difficult problems and limited resources), rather than to a generic weakness in the strategies used. The way that a problem presents itself to practitioners may make it very easy to entertain plausible but in fact erroneous possibilities.

Fixation may represent the downside of normally efficient and reliable cognitive processes given the cognitive demands of dynamic situations and cases where diagnosis and response happen in parallel (Woods, 1994). It is clear that in demanding situations where the state of the monitored process is changing rapidly, there is a potential conflict or tradeoff between the need to revise the situation assessment and the need to maintain coherence. Not every change is important; not every signal is meaningful. The practitioner whose attention is constantly shifting from one item to another may not be able to formulate a complete and coherent picture of the state of the system (thematic vagabonding in Dorner, 1983). Conversely, the practitioner whose attention does not shift may miss cues and data that are critical to updating the situation assessment. This latter condition may lead to fixation.

Given the kinds of cognitive processes that seem to be involved in fixation, there are a variety of techniques that, in principle, may reduce this form of breakdown. Data on successful and unsuccessful revision of erroneous situation assessments show that it usually takes a person with a fresh point of view on the situation to break a team or individual out of a fixation (Woods et al., 1987). Note that this result reveals a distributed, multi-agent component to cognition at work. Thus, one can change the architecture of the distributed system to try to ensure a fresh point of view, i.e., one that is unbiased by the immediate context. Another approach is to try to develop distributed system architectures where one person or group criticizes the assessments developed by the remainder of the group (e.g., a devil's advocate team member as in Schwenk and Cosier, 1980). A third direction is predicated on the fact that poor feedback about the state and behavior of the monitored process, especially related to goal achievement, is often implicated in fixations and failures to revise. Thus, one can provide practitioners with new kinds of representations about what is going on in the monitored process (e.g., Woods et al., 1987 for examples from nuclear power which tried this in response to the Three Mile Island accident).

Interacting Goals

Another set of factors that effect cognition at work is strategic in nature. People have to make tradeoffs between different but interacting or conflicting goals, between values or costs placed on different possible outcomes or courses of action, or between the risks of different errors. They must make these tradeoffs while facing irreducible uncertainty, risk, and the pressure of limited resources (e.g., time pressure; opportunity costs). One may think of these tradeoffs in terms of simplistic global examples like safety versus economy. Tradeoffs also occur on other kinds of dimensions. In responding to an anomaly in domains like aircraft or space vehicles, for example, there is a tradeoff with respect to when to commit to a course of action. Practitioners have to decide whether to take corrective action early in the course of an incident with limited information, to delay the response and wait for more data to come in, to search for additional findings, or to ponder additional alternative hypotheses.

Practitioners also trade off between following operational rules or taking action based on reasoning about the case itself (cf., Woods et al., 1987). Do the standard rules apply to this particular situation when some additional factor is present that complicates the textbook scenario? Should we adapt the standard plans or should we stick with them regardless of the special circumstances? Strategic tradeoffs can also involve coordination among people and machine agents in the distributed human-machine cognitive system (Woods et al., 1994 chapter 4). A machine expert recommends a particular diagnosis or action, but what if your own evaluation is different? What is enough evidence that the machine is wrong to justify disregarding the machine expert's evaluation and proceeding on your own evaluation of the situation?

Criterion setting on these different tradeoffs may not always be a conscious process or an explicit decision made by individuals. The criterion adopted may be an emergent property of systems of people, either small groups or larger organizations. The criterion may be fairly labile and susceptible to influence or relatively stable and difficult to change. The tradeoffs may create explicit choice points for practitioners embedded in an evolving situation, or may influence performance indirectly for example by shifting a team's mindset in a particular direction (e.g., Layton, Smith and McCoy, 1994).

Goal Conflicts

Multiple goals are simultaneously relevant in actual fields of practice. Depending on the particular circumstances in effect in a particular situation, the means to influence these multiple goals will interact, potentially producing conflicts between different goals. Expertise consists, in part, of being able to negotiate among interacting goals by selecting or constructing the means to satisfy all sufficiently or by deciding which goals to focus on and which goals to relax or sacrifice in a particular context.

However, practitioners may fail to meet this cognitive demand adequately. An adequate analysis of human performance and the potential for error requires explicit description of the interacting goals, the tradeoffs being made, and the pressures present that shift the operating points for these tradeoffs (Cook and Woods, 1994).

The finding potential conflicts, assessing their impact, and developing robust strategies may be quite difficult. Consider the anesthesiologist. Practitioners' highest level goal (and the one most often explicitly acknowledged) is to protect patient safety. But that is not the only goal. There are other goals, some of which are less explicitly articulated. These goals include reducing costs, avoiding actions that would increase the likelihood of being sued, maintaining good relations with the surgical service, maintaining resource elasticity to allow for handling unexpected emergencies, and others.

In a given circumstance, the relationships between these goals can produce conflicts. In the daily routine, for example, maximizing patient safety and avoiding lawsuits create the need to maximize information about the patient through preoperative workup. The anesthetist may find some hint of a potentially problematic condition and consider further tests that may incur costs, risks to the patient, and a delay of surgery. The cost-reduction goal provides an incentive for a minimal preoperative workup and the use of same-day surgery. This conflicts with the other goals. The anesthetist may be squeezed in this conflict—gathering the additional information, which in the end may not reveal anything important, will cause a delay of surgery and decrease throughput.

Given these external pressures some medical practitioners will not follow up hints about some aspect of the patient's history because to do so would impact the usual practices relative to throughput and economic goals. However, failing to acquire the information may reduce the ill-defined margin of safety and, in a specific case, the omission may turn out to be a contributor to an incident or accident. Other practitioners will adopt a conservative stance and order tests for minor indications even though the yield is low. This may affect the day's surgical schedule, the hospital and the surgeons' economic goals, and the anesthesiologists' relationship with the surgeons. In either case, the nature of the goals and pressures on the practitioner are seldom made explicit and rarely examined critically.

Analyses of past disasters frequently find that goal conflicts played a role in the accident evolution, especially when they place practitioners in double binds. In one tragic aviation disaster, the Dryden Ontario crash (Moshansky, 1992), several different organizational pressures to meet economic goals, along with organizational decisions to reduce resources, created a situation which placed a pilot in a double bind. Demand for the flight created a situation where the flight either could carry all of the passengers awaiting transport or carry enough fuel to reach their destination, but not both (no other aircraft was available to meet demand). The Captain decided to offload passengers and make a direct trip; the Company (a regional carrier) overruled him deciding to offload fuel and make a refueling stop at Dryden on the way to their destination.

After landing at Dryden weather was within nominal standards but a freezing rain had begun. Any delays increased the probability of wing ice. After landing, the crew had to keep engines running (a "hot" fuel) or they would have been unable to restart them (a) because the aircraft had departed with an inoperative auxiliary power unit and (b) due to the lack of equipment to restart jet engines at this airport. However, deicing is not permitted with an engine running because of other safety concerns. Several factors led to delays (e.g., they delayed on the ramp waiting for another aircraft to land). The captain attempted to takeoff, and the aircraft crashed due to wings contaminated with ice.

The immediate reaction was that a pilot error was responsible. In hindsight, there were obvious threats which it appeared the pilot recklessly disregarded. However, this view drastically oversimplified the situation the pilot faced. The actual accident investigation went deeper and found that organizational factors and specific circumstances placed the pilot in a goal conflict and double bind (note the report needed four volumes to lay out all of the different organizational factors and how they created the latent conditions for this accident).

One critical part of the investigation was understanding the situation from the point of view of the practitioners in the situation. When the investigators reconstructed the view of the evolving situation from the vantage point of a pilot, the Captain’s dilemmas became clear. Deciding not to take off would strand a full plane of passengers, disrupt schedules, and lose money for the carrier. Such a decision would be regarded as an economic failure with potential sanctions for the pilot. On the other hand, the means to accommodate the interacting constraints of the weather threat and refueling process were not available due to organizational choices not to invest in or cut back on equipment at peripheral airports such as Dryden. The air carrier was new to jet service and provided minimal support for the flight crew in terms of guidance or from company dispatch.

The accident investigation identified multiple, latent organizational factors which created the dilemmas. These included factors at the level of the regional carrier (e.g., new to jet operations but with inadequate investment in expertise and infrastructure to support these operations), at the level of the relationship between the regional carrier and the parent carrier (e.g., the parent organization distanced itself from operations and safety issues in its newly acquired subsidiary; the regional carrier minimized communication to preserve its autonomy), and at the level of regulatory oversight (e.g., breakdowns due to increased workload in a deregulated environment coupled with major staff and budget cuts).

As in this case, constraints imposed by organizational or social context can create or exacerbate competition between goals. Organizational pressures generated competition between goals that was an important factor in the breakdown of safety barriers in the system for transporting oil through Prince William Sound that preceded the *Exxon Valdez* disaster (National Transportation Safety Board, 1990).

It should not be thought that the organizational goals are necessarily simply the written policies and procedures of the institution. Indeed, the messages received by practitioners about the nature of the institution's goals may be quite different from those that management acknowledges. How the organization reacts to incidents, near misses and failures can send sharp end practitioners potent if implicit messages about how to tradeoff conflicting goals (Rochlin et al., 1987). These covert factors are especially insidious because they affect behavior and yet are unacknowledged. For example, the Navy sent an implicit but very clear message to its commanders by the differential treatment it accorded to the commander of the *Stark* following that incident (U.S. House of Representatives Committee on Armed Services, 1987) as opposed to the *Vincennes* following that incident (U.S. Department of Defense, 1988).

Goal conflicts also can arise from intrinsic characteristics of the field of activity. An example from cardiac anesthesiology is the conflict between the desirability of a high blood pressure to improve cardiac perfusion (oxygen supply to the heart muscle) and a low one to reduce cardiac work. The blood pressure target adopted by the anesthetist to balance this tradeoff depends in part on the practitioner's strategy, the nature of the patient, the kind of surgical procedure, anticipating potential risks (e.g., the risk of major bleeding), and the negotiations between different people in the operating room team (e.g., the surgeon who would like the blood pressure kept low to limit the blood loss at the surgical site).

Because local rationality revolves around how people pursue their goals, understanding performance at the sharp depends on tracing interacting multiple goals and how they produce tradeoffs, dilemmas, and double binds. In this process the investigator needs to understand how the tradeoffs produced by interacting goals are usually resolved. Practitioners in a field of activity may usually apply standard routines without deliberating on the nature of a dilemma, or they may work explicitly to find ways to balance the competing demands. The typical strategies for resolving tradeoffs, dilemmas and interacting goals, whether implicit in organizational practices or explicitly developed by practitioners, may be

- robust -- work well across a wide range of circumstance but still not guarantee a successful outcome,
- brittle -- work well under a limited set of conditions but breakdown when circumstances create situations outside the boundary conditions,
- poor -- very vulnerable to breakdown.

Understanding where and how goal conflicts can arise is a first step. Organizations can then examine how to improve strategies for handling these conflicts, even though ultimately there may be no algorithmic methods that can guarantee successful outcomes in all cases.

Problem Demands

In the above discussion, one should notice that it is difficult to examine cognition at work without also speaking of the demands that situations place on individuals, teams, and more distributed, coordinated activity (Rasmussen, 1986). The concept of local rationality captures this idea that cognitive activity needs to be considered in light of the <u>demands</u> placed on practitioners by characteristics of the incidents and problems that occur. These problem demands vary in type and degree. One incident may present itself as a textbook version of a well practiced plan while another may occur accompanied by several complicating factors which together create a more substantive challenge to practitioners. One case may be straightforward to diagnose, while another may represent a garden path problem (e.g., Johnson et al., 1988; Roth et. al., 1992). One case may be straightforward because a single response will simultaneously satisfy the multiple relevant goals, while another may present a dilemma because multiple important goals conflict.

Understanding the kinds of problem demands that can arise in a field of activity can reveal a great deal about the knowledge activation, attentional control or handling of multiple goals that is needed for successful performance. In other words, problem demands shape the cognitive activities of any agent or set of agents who might confront that incident. This perspective implies that one must consider what features of domain incidents and situations increase problem demands (coupling, escalation, time pressure, sources of variability particularly unanticipated variability, variations in tempo including rhythms of self-paced and event-driven activity, conflicting goals, uncertainty). The expression of expertise and error, then, is governed by the interplay of problem demands inherent in the field of activity and the resources available to bring knowledge to bear in pursuit of the critical goals. Analyses of the potential for failure look for mismatches in this demand-resource relationship (Rasmussen, 1986). Two examples of demand factors are coupling and escalation.

Coupling

One demand factor that affects the kinds of problems that arise is the degree of coupling in the underlying system (Perrow, 1984). Increasing the coupling in a system, increases the physical and functional interconnections between parts (Rasmussen, 1986). This has a variety of consequences but among them are effects at a distance, cascades of disturbances, side effects of actions (Woods, 1994). Several results follow from the fact that apparently distant parts are coupled.

- Practitioners must master new knowledge demands, e.g., knowing how different parts of the system interact physically or functionally. Greater investments will be required to avoid buggy, incomplete or oversimplified models. However, many of these interconnections will be relevant only under special circumstances. The potential for inert knowledge may go up.
- An action or a fault may influence a part of the system that seems distant from the area of focus. This complicates the diagnostic search process, for example, the ability to discriminate red herrings from important “distant” indications will be more difficult (e.g., the problem in Roth et al., 1992). In hindsight the relevance of such distant indications will be crystal clear, while from a practitioners’ perspective the demand is to discriminate irrelevant data from the critical findings to focus on.
- A single action will have multiple effects. Some of these will be the intended or main effects while others will be “side” effects. Missing side effects in diagnosis, planning and adapting plans in progress to cope with new events is a very common form of failure in highly coupled systems.

- A fault will produce multiple disturbances and these disturbances will cascade along the lines of physical and functional interconnection in the underlying monitored process. This also complicates diagnostic search (evidence will come in over time), but, importantly, it will increase the tempo of operations.

Escalation

A related demand factor is phenomenon of escalation which captures a fundamental correlation between the situation, the cognitive demands, and the penalties for poor human-machine interaction (Woods et al., 1994): the greater the trouble in the underlying process or the higher the tempo of operations, the greater the information processing activities required to cope with the trouble or pace of activities.

As situations diverge from routine or textbook, the tempo of operations escalates and the cognitive and cooperative work needed to cope with the anomalous situation escalates as well. As a result, demands for monitoring, attentional control, information gathering, and communication among team members (including human-machine communication) all tend to go up with the unusualness, tempo and criticality of situations. More knowledge and more specialist knowledge will need to be brought to bear as an incident evolves and escalates. This occurs through the technological and organizational structures used to access knowledge stored in different systems, places, or people. More lines of activity and thought will arise and need to be coordinated (Woods, 1994). The potential for goals to interact and conflict will go up. Focusing on the most critical goals may be needed to make tradeoff decisions. Plans will need to be put into effect to cope with the anomalous situation, but some practitioners will need to evaluate how well contingency plans match the actual circumstances or how to adapt them to the specific context.

If there are workload or other burdens associated with using a computer interface or with interacting with an autonomous or intelligent machine agent, these burdens tend to be concentrated at the very times when the practitioner can least afford new tasks, new memory demands, or diversions of his or her attention away from the job at hand to the interface per se (e.g., the case of cockpit automation; Billings, 1996). This means the penalties for poor design will be visible only in beyond-textbook situations. Since the system will seem to function well much of the time, breakdowns in higher tempo situations will seem mysterious after-the-fact and attributed to “pilot error” (see Sarter and Woods, in press, for this process in the case of cockpit automation).

Cognition in Context

While many of the factors discussed in the previous sections appear to be aspects of an individual's cognitive processes, i.e., knowledge organization, mental model, attention, judgment under uncertainty, we quickly find in work environments that these cognitive activities occur in the context of other practitioners, supported by artifacts of various types, and framed by the demands of the organization (Moray, this volume). While research on errors often began by considering the cognition of individuals, investigations quickly revealed one had to adopt a broader focus. For example, the more researchers have looked at success and failure in complex work settings, the more they have realized that a critical part of the story is how resources and constraints provided by the blunt end shape and influence the behavior of the people at the sharp end (Rochlin et al., 1987; Hirschhorn, 1993; Reason, 1998).

When investigators studied cognition at work, whether on flightdecks of commercial jet airliners, in control centers that manage space missions, in surgical operating rooms, in control rooms that manage chemical or energy processes, in control centers that monitor telecommunication networks, or in many other fields of human activity, first, they did not find cognitive activity isolated in a single individual, but rather they saw cognitive activity distributed across multiple agents (Resnick, Leavine, and Teasley, 1991; Hutchins, 1995). Second, they did not see cognitive activity separated in a thoughtful individual, but rather as a part of a stream of activity (Klein, Orasanu, and Calderwood, 1993). Third, they saw these sets of active practitioners embedded in larger group, professional, organizational, or institutional contexts which constrain their activities, set up rewards and punishments, define goals which are not always consistent, and provide resources (e.g., Hutchins, 1990). Moments of individual cognition punctuated this larger flow, and they were set up and conditioned by the larger system and communities of practice in which that individual was embedded.

Fourth, they observed phases of activity with transitions and evolution. Cognitive and physical activity varied in tempo, with periods of lower activity and more self-paced tasks interspersed with busy, externally paced operations where task performance was more critical. These higher tempo situations created greater need for cognitive work and at the same time often created greater constraints on cognitive activity (e.g., time pressure, uncertainty, exceptional circumstances, failures, and their associated hazards). They observed that there are multiple consequences at stake for the individuals, groups, and organizations involved in the field of activity or affected by that field of activity—such as economic, personal, and safety goals.

Fifth, they noticed that tools of all types are everywhere. Almost all activity was aided by external artifacts, some fashioned from traditional technologies and many built in the computer medium (Norman, 1993). More in-depth observation revealed that the computer technology was often poorly adapted to the needs of the practitioner. The computer technology was used clumsy in that the computer-based systems made new demands on the practitioner, demands that tended to congregate at the higher tempo or higher criticality periods. Close observation revealed that people and systems of people (operators, designers, regulators, etc.) adapted their tools and their activities continuously to respond to indications of trouble or to meet new demands. Furthermore, new machines were not used as the designers intended, but were shaped by practitioners to the contingencies of the field of activity in a locally pragmatic way (Woods et al., 1994, chapter 5).

These kinds of observations about cognition at work has led some to propose that cognition can be seen as fundamentally public and shared, distributed across agents, distributed between external artifacts and internal strategies, embedded in a larger context that partially governs the meanings that are made out of events (Winograd and Flores, 1988; Norman, 1993; Hutchins, 1995). Understanding cognition then depends as much on studying the context in which cognition is embedded and the larger distributed system of artifacts and multiple agents, as on studying what goes on between the ears -- a distributed cognitive system perspective (Hutchins, 1995).

Hughes Randall and Shapiro (1992, p. 5) illustrate this distributed cognitive system viewpoint in their studies of the UK air traffic control system and the reliability of this system.

> If one looks to see what constitutes this reliability, it cannot be found in any single element of the system. It is certainly not to be found in the equipment . . . for a period of several months during our field work it was failing regularly. . . . Nor is it to be found in the rules and procedures, which are a resource for safe operation but which can never cover every circumstance and condition. Nor is it to be found in the personnel who, though very highly skilled, motivated, and dedicated, are as prone as people everywhere to human error. Rather we believe it is to be found in the cooperative activities of controllers across the 'totality' of the system, and in particular in the way that it enforces the active engagement of controllers, chiefs, and assistants with the material they are using and with each other.

Success and failure belong to the larger operational system and not simply to an individual. Failure involves breakdowns in cognitive activities which are distributed across multiple practitioners and influenced by the artifacts used by those practitioners. This is perhaps best illustrated in processes of error detection and recovery which play a key role in determining system reliability in practice. In most domains, recovery involves the interaction of multiple people cross stimulating and cross checking each other (e.g., see Billings, 1996 for aviation; Rochlin et al., 1987 for air carrier operations; also see Seifert and Hutchins, 1992).

Hindsight and Cognitive Factors

Now that we have explored some of the factors that affect human performance, the way in which our knowledge of outcome biases attributions about causes of accidents becomes clearer. With knowledge of outcome, reviewers oversimplify the situation faced by practitioners in context (Woods et al., 1994). First, reviewers, after-the-fact assume that if people demonstrate knowledge in some context, then that knowledge should be available in all contexts. However, research on human performance shows that calling to mind knowledge is a significant cognitive process. Education research has focused extensively on the problem of inert knowledge—knowledge that can be demonstrated in one context (e.g., test exercises) is not activated in other contexts where it is relevant (e.g., ill-structured problems). Inert knowledge consists of isolated facts that are disconnected from how the knowledge can be used to accomplish some purpose. This research emphasizes the need to "conditionalize" knowledge to its use in different contexts as a fundamental training strategy.

Second, reviewers, after-the-fact, assume that if data are physically available, then their significance should be appreciated in all contexts. Many accident reports have a statement to the effect that all of the data relevant in hindsight were physically available to the people at the sharp end, but the people did not find and interpret the right data at the right time. Hindsight biases reviewers, blocking from view all of the processes associated with forming and shifting mindset as new events occur, the attentional demands of the situation, and the factors that led to a breakdown in synchronizing mindset in a changing world. Instead of seeing the mindset related factors and how they shape human performance, reviewers in hindsight are baffled by the practitioners inability to see what is obvious to all after-the-fact. In this explanatory vacuum reviewers in hindsight fall back on other explanations, for example, a motivational or effort factor (the people involved in an incident didn't "try hard enough") and try to improve things by a carrot and stick approach of exhortations to try harder combined with punishments for failure).

Third, since reviewers after-the-fact know which goal was the most critical, they assume that this should have been obvious to the practitioners working before-the-fact. However, this ignores the multiple interacting goals that are always present in systems under resource pressure, goals that co-exist sometimes peacefully and sometimes work against each other. If practitioner actions that are shaped by a goal conflict contribute to a bad outcome in a specific case, then it is easy for post-incident evaluations to say that a "human error" occurred—e.g., the practitioners should have delayed the surgical procedure to investigate the hint. The role of the goal conflict may never be noted; therefore, post-accident changes cannot address critical contributors, factors that can contribute to other incidents.

To evaluate the behavior of the practitioners involved in an incident, it is important to elucidate the relevant goals, the interactions among these goals, and the factors that influence how practitioners make tradeoffs in particular situations. The role of these factors is often missed in evaluations of the behavior of practitioners. As a result, it is easy for organizations to produce what appear to be solutions that in fact exacerbate conflict between goals rather than help practitioners handle goal conflicts in context. In part, this occurs because it is difficult for many organizations (particularly in regulated industries) to admit that goal conflicts and tradeoff decisions exist. However distasteful to admit or whatever public relations problems it creates, denying the existence of goal interactions does not make such conflicts disappear and is likely to make them even tougher to handle when they are relevant to a particular incident.

Debiasing and Studying Human Error

Hindsight biases fundamentally undermine our ability to understand the factors that influenced practitioner behavior. Given knowledge of outcome, reviewers will tend to <u>oversimplify</u> the problem-solving situation that was actually faced by the practitioner. The dilemmas, the uncertainties, the tradeoffs, the attentional demands, and double binds faced by practitioners may be missed or under-emphasized when an incident is viewed in hindsight. Typically, hindsight biases make it seem that participants failed to account for information or conditions that “should have been obvious” or behaved in ways that were inconsistent with the (now known to be) significant information. Possessing knowledge of the outcome, because of hindsight biases, trivializes the situation confronting the practitioner, who cannot know the outcome before-the-fact, and makes the “correct” choice seem crystal clear.

Because hindsight biases mask the real dilemmas, uncertainties, and demands practitioners confront, the result is a distorted view of the factors contributing to the incident or accident. In this vacuum, we can only see human performance after an accident or near miss as irrational, willing disregard (for what is now obvious to us and to them), or even diabolical. This seems to support the belief that human error often is the cause of an accident and that this judgment provides a satisfactory closure to the accident. When the label human error ends the investigation, the only response left is to search for the culprits and once they have been identified, remove them from practice, provide remedial training, replace them with technology, while new policies and procedures can be issued to keep other practitioners in line.

The difference between these everyday or "folk" reactions to failure and investigations of the factors that influence human performance is that researchers see the label human error as the starting point for investigations and use methods designed to remove hindsight biases to see better the factors that influenced the behavior of the people in the situation. When this is done, from Bessel's observations to Fitts' studies to today's investigations of computerized cockpits and other domains, the results identify the multiple deeper factors that lead to erroneous actions and assessments. These deeper factors will still be present even if the people associated with the failure are removed from practice. The vulnerabilities remain despite injunctions for practitioners to be more careful in the future. The injunctions or punishments that follow the accident may even exacerbate some of the vulnerabilities resident in the larger system.

We always can look back at people, episodes or cases in any system and using one or another standard identify any number of "errors," that is, violations of that standard. The key to safety is not minimizing or eradicating an infection of error (Rasmussen, 1990). Effective, robust, "high reliability" systems are able to recognize trouble before negative consequences occur. This means that processes involved in detecting that a situation is heading towards trouble and re-directing the situation away from a poor outcome is an important part of human performance related to safety versus failure. Evidence of difficulties, problems, incidents is an important form of information about the organization and operational system that is necessary for adaptive and constructive change (Reason, 1998). Studying cognitive factors, coordinative activities, and organizational constraints relative to problems demands in a particular domain is a crucial part of generating this base of information. Successful, "high reliability" organizations value, encourage, and generate such flows of information without waiting for accidents to occur (Rochlin, LaPorte, and Roberts, 1987).

Summary

Considering the topic "human error" quickly led us to two different perspectives.

"Human error" in one sense is a label invoked by stakeholders after-the-fact in a psychological and social process of causal attribution. Psychologists and others have studied these processes and in particular found a pernicious influence of knowledge of outcome, such as outcome and hindsight biases, that obscures the factors that shape human performance. Studying "human error" from this perspective is the study of how stakeholders react to failure.

"Human error" in another sense refers to the processes that lead up to success and failure (or potential failures) of a *system*. Researchers, using techniques to escape the hindsight bias, have studied such work domains and the factors that influence the performance of the people who work in those systems. In this chapter, we have reviewed a portion of the results using one heuristic structure -- knowledge in context, mindset and goal conflicts.

However, there is an irony in these research results. In focusing on cognitive factors behind the label human error, processes that would seem to reside only within individuals, researchers instead found

- a critical role for understanding the demands posed by problems, e.g., coupling and escalation (Perrow, 1984; Rasmussen, 1986),
- that artifacts shape cognition (Zhang and Norman, 1994; Woods et al., 1994, chapter 5),
- that coordination across practitioners is an essential component of success (Resnick, Levine and Teasley, 1991; Hutchins, 1995),
- that organizations create or exacerbate constraints on sharp end practitioners (Rochlin et al., 1987; Reason, 1998).

When one investigates the factors that produce “human error” in actual work environments, our view of cognition expands. “Cognition in the wild,” as Hutchins (1995) puts it, is distributed across people, across people and machines, and constrained by context.

References

Baron, J. and Hershey, J. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology, 54,* 569-579.

Boring, E.G. (1950). *A History of Experimental Psychology (2nd Edition)*. New York: Appleton-Century-Crofts.

Billings, C. E. (1996). *Aviation Automation: The Search For A Human-Centered Approach*. Hillsdale, NJ: Lawrence Erlbaum.

Bransford, J., Sherwood, R., Vye, N., and Rieser, J., (1986). Teaching and problem solving: Research foundations. *American Psychologist, 41,* 1078-1089.

Byrne, M. D. and Bovair, S. (1997). A working memory model of a common procedural error. *Cognitive Science*, 21, 31-61.

Caplan, R., Posner, K., and Cheney, F. (1991). Effect of outcome on physician judgments of appropriateness of care. *Journal of the American Medical Association, 265*, 1957-1960.

Chi, M. T. H., Glaser, R. and Farr, M. (1988). *The Nature Of Expertise.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Cook, R. I. and Woods, D. D. (1994). Operating at the 'Sharp End:' The Complexity of Human Error. In M.S. Bogner, editor, *Human Error in Medicine*, Hillsdale NJ: Lawrence Erlbaum.

De Keyser, V. and Woods, D. D. (1990). Fixation errors: Failures to revise situation assessment in dynamic and risky systems. In A. G. Colombo and A. Saiz de Bustamante (Eds.), *System Reliability Assessment*. The Netherlands: Kluwer Academic, 231-251.

Dorner, D. (1983). Heuristics and cognition in complex systems. In Groner, R., Groner, M. and Bischof, W. F. (Eds.), *Methods of Heuristics.* Hillsdale NJ: Lawrence Erlbaum.

Durso, F. T. and Gronlund, S. D. (this volume). Situation Awareness.

Feltovich, P. J., Ford, K. M. and Hoffman, R. R. (1997). *Expertise in Context*. Cambridge MA: MIT Press.

Feltovich, P. J., Spiro, R. J. and Coulson, R. (1989). The nature of conceptual understanding in biomedicine: The deep structure of complex ideas and the development of misconceptions. In D. Evans and V. Patel (Eds.), *Cognitive Science In Medicine: Biomedical Modeling.* Cambridge, MA: MIT Press.

Feltovich, P. J., Spiro, R. J., and Coulson, R. (1993). Learning, teaching and testing for complex conceptual understanding. In N. Fredericksen, R. Mislevy and I. Bejar (Eds.), *Test Theory For A New Generation Of Tests*., Hillsdale NJ: Lawrence Erlbaum.

Feltovich, P. J., Spiro, R. J., and Coulson, R. (1997). Issues of expert flexibility in contexts characterized by complexity and change. In Feltovich, P. J., Ford, K. M. and Hoffman, R. R. (Eds.), *Expertise in Context*. Cambridge MA: MIT Press.

Fischer, U., Orasanu, J.M. and Montvalo, M. (1993). Efficient decision strategies on the flight deck. In *Proceedings of the Seventh International Symposium on Aviation Psychology*. Columbus, OH. April.

Fischhoff, B. (1975). Hindsight ≠ foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 288-299.

Fischhoff, B. (1982). For those condemned to study the past: Heuristics and biases in hindsight. In D. Kahneman, P. Slovic and A. Tversky (Eds.), *Judgment under Uncertainty: Heuristics and biases.* Cambridge, MA: Cambridge University Press

Fitts, P. M, and Jones, R. E. (1947). *Analysis of factors contributing to 460 "pilot-error" experiences in operating aircraft controls* (Memorandum Report TSEAA-694-12). Wright Field, OH: U.S. Air Force Air Materiel Command, Aero Medical Laboratory.

Gaba, D. M., and DeAnda, A. (1989). The response of anesthesia trainees to simulated critical incidents. *Anesthesia and Analgesia, 68*, 444-451.

Gentner, D., and Stevens, A. L. (Eds.) (1983). *Mental Models*. Hillsdale NJ: Lawrence Erlbaum Associates.

Getty, D. J., Swets, J. A., Pickett, R. M. and Gonthier, D. (1995) System operator response to warnings of danger: A laboratory invesitgation of the effects of predictive value of a warning on human response time. *Journal of Experimental Psychology: Applied*, 1, 19-33.

Gilson, R. D. (Ed.) (1995). Special Section on Situation Awareness. *Human Factors*, 37, p. 3-157.

Gopher, D. (1991). The skill of attention control: Acquisition and execution of attention strategies. In *Attention and Performance XIV*. Hillsdale, NJ: Lawrence Erlbaum.

Hilton, D. (1990). Conversational processes and causal explanation. *Psychological Bulletin, 197*, 65-81.

Hirschhorn, L. (1993). Hierarchy vs. bureaucracy: The case of a nuclear reactor. In K. H. Roberts (Ed.), *New Challenges To Understanding Organizations*. New York: McMillan.

Hollnagel. E. (1993). *Human Reliability Analysis: Context and Control.* London: Academic Press.

Hughes, J., Randall, D. and Shapiro, D. (1992). Faltering from ethnography to design. *Computer Supported Cooperative Work (CSCW) Proceedings. (November)*, 1-8.

Hutchins, E. (1990). The technology of team navigation. In J. Galegher, R. Kraut, and C. Egido (Eds.), *Intellectual Teamwork: Social and technical bases of cooperative work*. Hillsdale, NJ: Lawrence Erlbaum.

Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.

Jager Adams, M., Tenney, Y. J., and Pew, R. W. (1995). Situation awareness and the cognitive management of complex systems. *Human Factors*, 37, 85-104.

Johnson, P. E., Duran, A. S., Hassebrock, F., Moller, J., Prietula, M., Feltovich, P. J., Swanson, D. B. (1981). Expertise and error in diagnostic reasoning. *Cognitive Science, 5,* 235-283.

Johnson, P. E., Moen, J. B. and Thompson, W. B. (1988). Garden path errors in diagnostic reasoning. In L. Bolec and M. J. Coombs (Eds.), *Expert System Applications*. New York: Springer-Verlag.

Johnson, P. E., Jamal, K. and Berryman, R. G. (1991). Effects of framing on auditor decisions. *Organizational Behavior and Human Decision Processes*, 50, 75-105.

Johnson, P. E., Grazioli, S., Jamal, K. and Zualkernan, I. (1992). Success and failure in expert reasoning. *Organizational Behavior and Human Decision Processes*, 53, 173-203.

Kelley, H. H. (1973). The process of causal attribution. *American Psychologist, 28,* 107-128.

Klein, G. A., Orasanu, J., and Calderwood, R. (Eds.) (1993). *Decision Making in Action: Models and Methods,* Norwood NJ: Ablex.

Klein, G. A. (1998). *Sources of Power: How People Make Decisions*. MIT Press, Cambrdige MA.

LaBerge, D. (1995). *Attentional Processing: The Brain's Art of Mindfulness*. Harvard University Press, Cambridge MA.

Layton, C., Smith, P., and McCoy, E. (1994). Design of a cooperative problem-solving system for enroute flight planning: An empirical evaluation. *Human Factors*, 36, 94-119.

Lipshitz, R. (1989). "Either a medal or a corporal:" The effects of success and failure on the evaluation of decision making and decision makers. *Organizational Behavior and Human Decision Processes, 44,* 380-395.

Moray, N. (this volume). The Cognitive Psychology of Industrial Systems: An Introduction to Cognitive Engineering.

Moshansky, V. P. (1992). *Final report of the commission of inquiry into the Air Ontario crash at Dryden, Ontario*. Ottawa: Minister of Supply and Services, Canada.

National Transportation Safety Board. (1990). *Marine accident report: Grounding of the U.S. Tankship Exxon Valdez on Bligh Reef, Prince William Sound, near Valdez, Alaska, March 24, 1989. Report no. NTSB/MAT-90/04.* Springfield, VA: National Technical Information Service.

Neisser, U. (1976). *Cognition and Reality: Principles and Implications of Cognitive Psychology.* San Francisco, CA: W. H. Freeman and Company.

Newell, A. (1982). The knowledge level. *Artificial Intelligence, 18,* 87-127.

Norman, D. A. (1981). Categorization of action slips. *Psychological Review, 88,* 1-15.

Norman, D. A. (1983). Design rules based on analysis of human error. *Communications of the ACM, 26,* 254-258.

Norman, D. A. (1988). *The Psychology of Everyday Things.* New York, NY: Basic Books.

Norman, D. A. (1993). *Things That Make Us Smart.* Reading, MA: Addison-Wesley.

Patterson, E.S., Watts-Perotti, J.C. and Woods, D.D. Voice Loops as Coordination Aids in Space Shuttle Mission Control. *Computer Supported Cooperative Work,* in press.

Payne, J. W. and Bettman, J. R. and Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14,* 534-552.

Payne J. W., Johnson E. J., Bettman J. R. and Coupey, E. (1990). Understanding contingent choice: A computer simulation approach. *IEEE Transactions on Systems, Man, and Cybernetics, 20*, 296-309.

Perkins, D., Martin, F. (1986). Fragile knowledge and neglected strategies in novice programmers. In E. Soloway and S. Iyengar (Eds.), *Empirical Studies of Programmers.* Norwood, NJ: Ablex.

Perrow, C. (1984). *Normal Accidents: Living with High Risk Technologies*. New York, NY: Basic Books.

Rasmussen, J. (1986). *Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering.* New York: North-Holland.

Rasmussen, J. (1990). The role of error in organizing behavior. *Ergonomics*, 33, 1185-1199.

Rasmussen, J., Duncan, K. and Leplat, J. (1987). *New Technology and Human Error*. Chichester, John Wiley & Sons.

Reason, J. and Mycielska, K. (1982). *Absent Minded? The Psychology of Mental Lapses and Everyday Errors*. Englewood Cliffs, NJ: Prentice Hall.

Reason, J. (1990). *Human Error.* Cambridge, England: Cambridge University Press.

Reason, J. (1997). *Managing the Risks of Organizational Accidents.* Brookfield VT: Ashgate Publishing.

Resnick, L., Levine, J and Teasley, S. D. (1991). *Perspectives on Socially Shared Cognition*. New York: American Psychological Association.

Roberts, K. H., and Rousseau, D. M. (1989). Research in nearly failure-free, high-reliability organizations: Having the bubble. *IEEE Transactions in Engineering Management, 36*, 132-139.

Rochlin, G., LaPorte, T. R. and Roberts, K. H. (1987). The self-designing high reliability organization: Aircraft carrier flight operations at sea. *Naval War College Review, (Autumn),* 76-90.

Roth E. M., Woods, D. D. and Pople, H. E., Jr. (1992). Cognitive simulation as a tool for cognitive task analysis. *Ergonomics*, *35*, 1163-1198.

Schwenk, C., and Cosier, R. (1980). Effects of the expert, devil's advocate and dialectical inquiry methods on prediction performance. *Organizational Behavior and Human Decision Processes. 26.*

Seifert, C. M. and Hutchins, E. (1992). Error as opportunity: Learning in a cooperative task. *Human-Computer-Interaction, 7*, 409-435.

Simon, H. (1957). *Models of Man (Social and Rational)*. New York: John Wiley and Sons.

Spiro, R. J., Coulson, R. L., Feltovich, P. J. and Anderson, D. K. (1988). Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains. *Proceedings of the Tenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Tasca, L. (1990). The Social Construction of Human Error. Unpublished Doctoral Dissertation, State University of New York at Stony Brook, August 1990.

U.S. Department of Defense. (1988). *Report of the formal investigation into the circumstances surrounding the downing of Iran Air Flight 655 on 3 July 1988.* Washington DC: Department of Defense.

U.S. House of Representatives Committee on Armed Services. (1987). *Report on the staff investigation into the Iraqi attack on the USS Stark.* Springfield, VA: National Technical Information Service.

Weick, K. E. and Roberts, K. H. (1993). Collective mind and organizational reliability: The case of flight operations on an aircraft carrier deck. *Administration Science Quarterly*, 38, 357--381.

Winograd, T. and Flores, F. (1987). *Understanding Computers and Cognition.* Reading, MA: Addison-Wesley.

Woods, D. D. (1994). Cognitive Demands and Activities in Dynamic Fault Management: Abduction and Disturbance Management. In N. Stanton, editor, *Human Factors of Alarm Design*, Taylor & Francis, London.

Woods, D. D. (1995). The alarm problem and directed attention in dynamic fault management. *Ergonomics*, 38(11), 2371-2393.

Woods, D. D. and Sarter, N. B. (in press). Learning from Automation Surprises and Going Sour Accidents. In N. Sarter and R. Amalberti (Eds.), *Cognitive Engineering in the Aviation Domain*, Erlbaum, Hillsdale NJ.

Woods, D. D. and J.C. Watts, J. C. How Not To Have To Navigate Through Too Many Displays. In Helander, M.G., Landauer, T.K. and Prabhu, P. (Eds.) *Handbook of Human-Computer Interaction*, 2nd edition. Amsterdam, The Netherlands: Elsevier Science, 1997.

Woods, D. D., Johannesen, L., Cook, R. I. and Sarter, N. B. (1994). *Behind Human Error: Cognitive Systems, Computers and Hindsight*. Crew Systems Ergonomic Information and Analysis Center, WPAFB, Dayton OH.

Woods, D. D., O'Brien, J. and Hanes, L. F. (1987). Human factors challenges in process control: The case of nuclear power plants. In G. Salvendy (Ed.), *Handbook of Human Factors/Ergonomics (First Edition).* New York: Wiley.

Yantis, S. (1993). Stimulus-drive attention capture. *Current Directions in Psychological Science*, 2, 156-161.

Zhang, J. and Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science*, 18, 87--122.

Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science*, 21, 179--217.